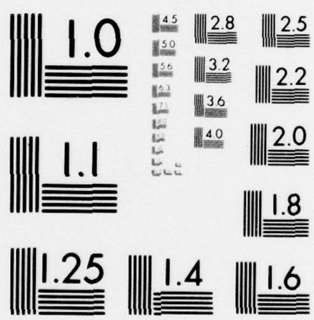
AD-A077 098 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF SCALED VARIABL--ETC(U)
JUN 79 K RITTER DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1967 NL

1 OF 1
AD
A077098

END
DATE
FILMED
12-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(9) SC

MRC Technical Summary Report # 1967

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF SCALED VARIABLE METRIC METHODS

Klaus Ritter

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

See 1473 in back

June 1979

(Received May 31, 1979)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF SCALED VARIABLE METRIC METHODS

Klaus Ritter

Technical Summary Report #1967

June 1979

ABSTRACT

This paper considers a class of variable metric methods for unconstrained minimization. The update formulas are such that the quasi-Newton equation is not necessarily satisfied. Under appropriate assumptions on the function to be minimized, each algorithm in this class converges globally and superlinearly. -> to p. -B-

AMS (MOS) Subject Classification 90C30.

Key Words: Unconstrained minimization, variable metric method, global convergence, superlinear convergence.

Work Unit Number 5 - Mathematical Programming and Operations Research.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

SIGNIFICANCE AND EXPLANATION

Cont.

Many practical problems in operations research may be reduced to minimizing a function with or without constraints. By means of penalty functions and similar techniques a constrained minimization problem can be converted into a sequence of unconstrained minimization problems. In this paper we discuss a class of algorithms for unconstrained minimization problems which converge rapidly to the solution from a starting point which is not necessarily a good approximation to the solution of the given problem.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

k. -B-

GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF SCALED VARIABLE METRIC METHODS

Klaus Ritter

1. Introduction

Variable metric methods are successfully used for iteratively calculating a sequence $\{x_j\}$ which converges rapidly to a global minimizer of a smooth convex function $F(x)$. At each point x_j the computation of the search direction is based on a quadratic model $Q(x)$ which interpolates the function value and the gradient of $F(x)$ at x_j as well as the gradient of $F(x)$ at x_{j-1}. In general $Q(x_{j-1}) \neq F(x_{j-1})$. The matrix defining the quadratic form in this model is updated in such a way that it satisfies the quasi-Newton equation.

In this paper a class of modified update formulas is considered which result in a quadratic model $Q(x)$ which in addition to interpolating the function value and the gradient of $F(x)$ at x_{j-1} minimizes the criterion

$$t \| v \|^2 + (1-t)(Q(x_{j-1}) - F(x_{j-1}))^2 .$$

Here v is the error in the quasi-Newton equation and t, $0 \leq t \leq 1$, is a weight factor which reflects the importance assigned to satisfying the quasi-Newton equation as opposed to interpolating $F(x_{j-1})$. Under the usual assumptions it is shown that the resulting variable metric methods converge globally and superlinearly.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

The so-called restricted Broyden methods [2] are contained in this new class of variable metric methods. They are obtained if the weight factor t is chosen equal to one.

In [1] Biggs describes a modification of the Broyden-Fletcher-Goldfarb-Shanno method which results in an update formula for which the quasi-Newton equation is also not satisfied. However, this method is not contained in our class. Recently, Schnabel [7] has shown that Biggs' method converges superlinearly.

2. Basic properties of scaled variable metric methods

Let $x \in E^n$ and let $F(x)$ be a real valued function. If $F(x)$ is twice differentiable at some point x_j we denote the gradient and the Hessian matrix of $F(x)$ at x_j by $\nabla F(x_j) = g_j$ and $G_j = G(x_j)$, respectively. A prime is used to denote the transpose of a vector or a matrix. For every $x \in E^n$, $\| x \|$ denotes the Euclidean norm of x.

Throughout this paper we make the following

Assumption 1

$F(x)$ is twice continuously differentiable and there are numbers $0 < \mu < \eta$ such that

$$\mu \| x \|^2 \leq x'G(y)x \leq \eta \| x \|^2 \quad \text{for all } x, y \in E^n .$$

It is well-known that Assumption 1 implies that $F(x)$ is uniformly convex and that there is a unique $z \in E^n$ such that

$$F(z) < F(x) \quad \text{for all} \quad x \in E^n \, , \quad x \neq z \, .$$

It is the purpose of this paper to discuss a class of algorithms which construct a sequence

$$\text{(2.1)} \qquad x_{j+1} = x_j - \sigma_j s_j \quad \text{with} \quad F(x_{j+1}) < F(x_j) \, , \quad j = 0,1,\ldots \, ,$$

which converges superlinearly to z. Here $s_j \in E^n$ is called a search direction and the scalar σ_j is referred to as the step size.

We say a sequence $\{x_j\}$ converges superlinearly to z if and only if

$$\frac{\| x_{j+1} - z \|}{\| x_j - z \|} \to 0 \quad \text{as} \quad j \to \infty \, .$$

Since z is a global minimizer of $F(x)$ we have $\nabla F(z) = 0$. Thus, if $x_j \to z$ as $j \to \infty$ then $g_j \to 0$ as $j \to \infty$. Furthermore, it follows from Lemma 1, proved at the end of this section, that $\{x_j\}$ converges superlinearly to z if and only if $\{g_j\}$ converges superlinearly to 0. Therefore, it suffices to show that

$$\text{(2.2)} \qquad \frac{\| g_{j+1} \|}{\| g_j \|} \to 0 \quad \text{as} \quad j \to \infty \, .$$

In order to derive conditions under which (2.2) holds, we assume for the moment that the sequence (2.1) converges to z. Then it follows from Taylor's theorem that, for every j,

$$\text{(2.3)} \qquad \begin{aligned} g_{j+1} &= g_j - \sigma_j \left(\int_0^1 G(x_j - t\sigma_j s_j) dt \right) s_j \\ &= g_j - \sigma_j G s_j - \sigma_j E_j s_j \, , \end{aligned}$$

where $G = G(z)$ and

$$(2.4) \qquad E_j = \int_0^1 G(x_j - t\sigma_j s_j)dt - G .$$

Dividing (2.3) by $\| g_j \|$ and using the triangle inequality we obtain the two inequalities

$$(2.5) \qquad \frac{\| g_{j+1} \|}{\| g_j \|} \leq \left\| \frac{g_j}{\| g_j \|} - G \frac{\sigma_j s_j}{\| g_j \|} \right\| + \left\| E_j \right\| \frac{\| \sigma_j s_j \|}{\| g_j \|}$$

and

$$(2.6) \qquad \left| \left\| \frac{g_j}{\| g_j \|} - G \frac{\sigma_j s_j}{\| g_j \|} \right\| - \left\| E_j \frac{\sigma_j s_j}{\| g_j \|} \right\| \right| \leq \frac{\| g_{j+1} \|}{\| g_j \|}$$

Since

$$\| E_j \| \leq \| \int_0^1 G(x_j - t\sigma_j s_j)dt - G(x_j) \| + \| G(x_j) - G \|$$

implies that

$$(2.7) \qquad \| E_j \| \to 0 \text{ as } j \to \infty$$

and it follows from Assumption 1 that

$$\| \sigma_j s_j \| = 0(\| g_j \|)$$

we deduce from (2.5), (2.6) and (2.7) that the sequence $\{g_j\}$ converges superlinearly to zero if and only if

$$\left\| \frac{g_j}{\| g_j \|} - G \frac{\sigma_j s_j}{\| g_j \|} \right\| \to 0 \text{ as } j \to \infty .$$

This shows that in order to achive superlinear convergence we have to determine the search direction s_j and the step size σ_j in such a way that

$$(2.8) \qquad \left\| \frac{g_j}{\| g_j \|} - \frac{Gs_j}{\| Gs_j \|} \right\| \to 0 \text{ as } j \to \infty .$$

and

$$(2.9) \qquad \left| \sigma_j \frac{\| Gs_j \|}{\| g_j \|} - 1 \right| \to 0 \text{ as } j \to \infty .$$

If a variable metric method is used to compute the sequence (2.1) then a matrix H_j is associated with each x_j and

$$(2.10) \qquad s_j = H_j g_j .$$

The matrix H_{j+1} is determined by adding a matrix of rank 1 or rank 2 to H_j in such a way that the quasi-Newton equation,

$$(2.11) \qquad H_{j+1} d_j = p_j ,$$

is satisfied. Here

$$(2.12) \qquad d_j = \frac{g_j - g_{j+1}}{\| \sigma_j s_j \|} , \quad p_j = \frac{s_j}{\| s_j \|} .$$

This procedure can be motivated as follows. If H_j approximates G^{-1} in the sense that

$$(2.13) \qquad \| H_j - G^{-1} \| \to 0 \quad \text{as} \quad j \to \infty ,$$

then it is clear that s_j as defined by (2.10) satisfies (2.8). Furthermore, it follows from (2.3) that

$$(2.14) \qquad G^{-1} d_j = p_j + G^{-1} E_j p_j .$$

Thus if H_{j+1} satisfies the quasi-Newton equation (2.11) we have

$$\| (H_{j+1} - G^{-1}) d_j \| \to 0 \quad \text{as} \quad j \to \infty \; .$$

It should, however, be observed that the condition (2.13) is not necessary. According to (2.8) we only need the property

$$\left\| \frac{G^{-1} g_j}{\| G^{-1} g_j \|} - \frac{H_j g_j}{\| H_j g_j \|} \right\| \to 0 \quad \text{as} \quad j \to \infty$$

combined with a step size σ_j satisfying (2.9).

Assuming that H_{j+1} is symmetric and positive definite we denote the inverse of H_{j+1} by B_{j+1} and consider the quadratic function

(2.15) $$Q(x) = F(x_{j+1}) + g'_{j+1}(x - x_{j+1}) + \frac{1}{2}(x - x_{j+1})' B_{j+1}(x - x_{j+1}) \; .$$

Obviously we have

(2.16) $$Q(x_{j+1}) = F(x_{j+1}) \quad \text{and} \quad \nabla Q(x_{j+1}) = g_{j+1} \; .$$

Furthermore, if H_{j+1} satisfies the quasi-Newton equation (2.11) then

(2.17) $$B_{j+1} p_j = d_j$$

and it follows that

(2.18) $$\nabla Q(x_j) = g_{j+1} + B_{j+1}(x_j - x_{j+1}) = g_j \; .$$

With $s_{j+1} = B^{-1}_{j+1} g_{j+1}$ it is easy to verify that $x_{j+1} - s_{j+1}$ is the unique global minimizer of the quadratic function (2.15). Since $B^{-1}_{j+1} = H_{j+1}$ we can interprete the negative search direction $-s_{j+1} = -H_{j+1} g_{j+1}$, associated by a variable metric method with a point x_{j+1}, as the direction which leads from x_{j+1} to the global minimizer of the quadratic function (2.15) which has the properties (2.16) and (2.18).

Since in general $Q(x_j) \neq F(x_j)$ we could try to determine the quadratic function $Q(x)$ such that

$$Q(x_j) = F(x_j) \tag{2.19}$$

by giving up one of the properties (2.16) or (2.18). Since at x_{j+1} the function value $F(x_{j+1})$ and the gradient g_{j+1} represent the most recent information about $F(x)$ that is available to us, it appears reasonable to insist on (2.16). However, then the quasi-Newton-equation (2.17) necessarily leads to the equality (2.18). Therefore, if we want to use a quadratic function of type (2.15) that has the properties (2.16) and (2.19) we have to relax the quasi-Newton equation (2.17).

Multiplying the equality (2.14) by G we obtain

$$Gp_j = d_j + E_j p_j \ .$$

Thus the quasi-Newton equation (2.17) guarantees that

$$\| B_{j+1} p_j - Gp_j \| \to 0 \quad \text{as} \quad j \to \infty \ .$$

However since Gp_j differs in general from d_j there is really no reason to insist on (2.17). On the contrary it might be better to relax (2.17) in favor of a better approximation of $F(x_j)$ by $Q(x_j)$.

In the following we will, therefore, consider a quadratic function (2.15) where B_{j+1} is symmetric positive definite and satisfies the relaxed quasi-Newton equation

$$B_{j+1} p_j = d_j + v_j \ . \tag{2.20}$$

We assume that v_j is an optimal solution to the problem

$$(2.21) \qquad \min_v \{ t \, \| v \|^2 + (1-t)(Q(x_j) - F(x_j))^2 \} .$$

Here the parameter t is restricted to the interval $0 \le t \le 1$. It can be interpreted as a weight factor which reflects the importance we assign to satisfying *the quasi-Newton equation* as opposed to interpolation $F(x_j)$.

Using (2.20) we obtain from (2.15)

$$(2.22) \quad Q(x_j) = F(x_{j+1}) + g'_{j+1}(x_j - x_{j+1}) + \frac{1}{2}(x_j - x_{j+1})'(g_j - g_{j+1}) + \frac{1}{2} \| x_j - x_{j+1} \| (x_j - x_{j+1})' v_j$$

$$= F(x_{j+1}) + \frac{1}{2}(g_j + g_{j+1})'(x_j - x_{j+1}) + \frac{1}{2} \| x_j - x_{j+1} \|^2 p_j' v_j .$$

Setting

$$(2.23) \qquad \nu_j = F(x_{j+1}) - F(x_j) + \frac{1}{2}(g_j + g_{j+1})'(x_j - x_{j+1})$$

we can therefore write (2.21) in the equivalent form

$$(2.24) \quad \min_v \{ t \| v \|^2 + (1-t)(\nu_j + \frac{1}{2} \| x_j - x_{j+1} \|^2 p_j' v)^2 \} .$$

Clearly v_j is an optimal soltuion to (2.24) if and only if

$$(2.25) \qquad t v_j + (1-t)(\nu_j + \frac{1}{2} \| x_j - x_{j+1} \|^2 p_j' v_j) p_j = 0 .$$

If $t > 0$ this implies $v_j = \lambda_j p_j$ for some λ_j . In order to obtain λ_j we substitute into (2.25) and solve for λ_j . This gives

$$\lambda_j = \frac{2(t-1)\nu_j}{2t + (1-t) \| x_j - x_{j+1} \|^2} .$$

Therefore,

$$(2.26) \qquad v_j = \frac{2(t-1)\nu_j}{2t + (1-t) \| x_j - x_{j+1} \|^2} p_j$$

is an optimal solution for (2.24). If $t > 0$, then (2.26) is the unique optimal solution. Furthermore, if $t = 1$, then $v_j = 0$ and (2.20) reduces to the quasi-Newton equation (2.17). On the other hand if $t = 0$, then

$$v_j = \frac{-2\nu_j}{\| x_j - x_{j+1} \|^2} p_j$$

and it follows from (2.22) and (2.23) that $Q(x_j) = F(x_j)$.

With

$$\delta_j = \delta_j(t) = \frac{2(t-1)\nu_j}{2t + (1-t)\|x_j - x_{j+1}\|^2} \tag{2.27}$$

we obtain from (2.20) and (2.26) the relaxed quasi-Newton equation

$$B_{j+1}p_j = d_j + \delta_j p_j ,$$

or in terms of H_{j+1},

$$H_{j+1}(d_j + \delta_j p_j) = p_j . \tag{2.28}$$

By Taylor's theorem there are vectors

$$z_j , \hat{z}_j \in \{ x \mid x = x_{j+1} + \lambda(x_j - x_{j+1}) , \; 0 \le \lambda \le 1 \} \tag{2.29}$$

such that

$$F(x_j) = F(x_{j+1}) + g'_{j+1}(x_j - x_{j+1}) + \frac{1}{2}(x_j - x_{j+1})'G(\hat{z}_j)(x_j - x_{j+1})$$

and

$$g'_j(x_j - x_{j+1}) = g'_{j+1}(x_j - x_{j+1}) + (x_j - x_{j+1})'G(z_j)(x_j - x_{j+1}) .$$

Therefore, it follows from (2.23) that

$$\nu_j = \frac{1}{2}(x_j - x_{j+1})'(G(z_j) - G(\hat{z}_j))(x_j - x_{j+1})$$

which by (2.27) implies that there is a constant δ such that

$$(2.30) \qquad |\delta_j(t)| \leq (1-t)\delta \|G(z_j) - G(\hat{z}_j)\| , \quad 0 \leq t \leq 1 .$$

Setting

$$\tilde{d}_j = d_j + \delta_j p_j$$

and choosing arbitrary parameters β_1 and β_2 with $\beta_1^2 + \beta_2^2 > 0$ we define the following class of update formulas

$$(2.31) \qquad H_{j+1} = H_j + \frac{\beta_1(\tilde{d}_j' p_j + \tilde{d}_j' H_j \tilde{d}_j) + \beta_2 \tilde{d}_j' H_j \tilde{d}_j}{(\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j)\tilde{d}_j' p_j} p_j p_j'$$

$$- \beta_1 \frac{p_j \tilde{d}_j' H_j + H_j \tilde{d}_j p_j'}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j} - \beta_2 \frac{H_j \tilde{d}_j \tilde{d}_j' H_j}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j} .$$

Assuming that all denominators in (2.31) are different from zero it is easy to verify that each matrix H_{j+1} defined by (2.31) satisfies the equation (2.28). Moreover, if H_j is symmetric so is H_{j+1} .

If we choose $t = 1$, i.e. $\tilde{d}_j = d_j$, then (2.31) is identical with a class of update formulas introduced by Broyden (see [2] and [3]).

In order to determine conditions on β_1 and β_2 which guarantee that the update formula (2.31) maintains the positive definiteness of H_j we define the two subspaces

$$S_j = \operatorname{span} \{ g_j, \tilde{d}_j \}$$

$$T_j = \{ x \mid (H_j g_j)'x = 0 , (H_j \tilde{d}_j)'x = 0 \}$$

and observe that

$$(2.32) \qquad H_{j+1}x = H_j x \quad \text{for} \quad x \in T_j .$$

Since $g_j \notin T_j$ and $\tilde{d}_j \notin T_j$ it follows that H_{j+1} is completely defined if we specify it on the subspace S_j. For each update formula (2.31) we have

$$(2.33) \qquad H_{j+1}\tilde{d}_j = p_j .$$

If g_j and d_j are linearly dependent, then H_{j+1} is completely defined by (2.32) and (2.33). If g_j and $\tilde{d}_j$ are linearly independent let $w_j \in S_j$ be such that

$$w_j' p_j = 0 \text{ and } H_j w_j = q_j \text{ with } \| q_j \| = 1 .$$

Then

$$\tilde{d}_j' H_{j+1} w_j = p_j' w_j = 0$$

and

$$H_{j+1} w_j = q_j - \frac{(\beta_1 p_j + \beta_2 H_j \tilde{d}_j)\tilde{d}_j' q_j}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j} \in \text{span}\{ q_j, p_j \} .$$

This implies that there is a vector $u_j \in \text{span}\{q_j, p_j\}$ such that

$$\| u_j \| = 1 , \quad \tilde{d}_j u_j = 0 , \quad w_j' u_j > 0$$

and, for every update formula (2.31),

$$(2.34) \qquad H_{j+1} w_j = \omega_j u_j ,$$

where only the parameter ω_j depends on the particular values of β_1 and β_2.

Since we assume that H_j is symmetric and positive definite it is not difficult to verify (see [6]) that H_j can be written in the form

$$(2.35) \qquad H_j = \frac{p_j p_j'}{\rho_j g_j' p_j} + \frac{q_j q_j'}{w_j' q_j} + \hat{H}_j$$

where $\rho_j = \| H_j g_j \|^{-1}$ and $\hat{H}_j$ is a matrix of rank $n-2$ with

$$\hat{H}_j g_j = \hat{H}_j w_j = 0 .$$

Now it follows from (2.32), (2.33), (2.34) and (2.35) that

$$(2.36) \qquad H_{j+1} = \frac{p_j p_j'}{\tilde{d}_j' p_j} + \omega_j \frac{u_j u_j'}{w_j' u_j} + \hat{H}_j .$$

Therefore, if H_j is positive definite, then H_{j+1} is positive definite if and only if

$$(2.37) \qquad \tilde{d}_j p_j > 0 \quad \text{and} \quad \omega_j > 0 .$$

Because $\tilde{d}_j p_j = d_j' p_j + \delta_j(t)$ and $d_j' p_j > \mu$ (see Lemma 1) we can force $\tilde{d}_j' p_j$ to be positive if we choose t sufficiently close to one (see (2.30)). In order to study the dependence of ω_j on β_1 and β_2 we first choose $\beta_1 = 1$ and $\beta_2 = 0$. If $t = 1$ the resulting update formula (2.31) corresponds to the Broyden-Fletcher-Goldfarb-Shanno-method [3], [4], [5], [8]. Then

$$H_{j+1} w_j = q_j - \frac{\tilde{d}' q_j}{\tilde{d}_j' p_j} p_j .$$

Thus, setting $\alpha_j = \tilde{d}' q_j / \tilde{d}_j' p_j$ we have

$$(2.38) \qquad u_j = \frac{q_j - \alpha_j p_j}{\| q_j - \alpha_j p_j \|} , \quad \omega_j = \| q_j - \alpha_j p_j \| .$$

Since by (2.35) ,

$$(2.39) \qquad H_j\tilde{d}_j = p_j \frac{\tilde{d}_j' p_j}{\rho_j g_j' p_j} + q_j \frac{\tilde{d}_j' q_j}{w_j' q_j}$$

we obtain for a general choice of the parameters β_1 and β_2

$$H_{j+1} w_j = \frac{q_j(\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j) - \beta_1 p_j \tilde{d}_j' q_j - \beta_2 H_j \tilde{d}_j \tilde{d}_j' q_j}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j}$$

$$= \frac{\beta_1 \tilde{d}_j' p_j + \beta_2 (\tilde{d}_j' p_j)^2 / \rho_j g_j' p_j}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j} (q_j - \alpha_j p_j) .$$

Thus in general

$$(2.40) \qquad \omega_j = \gamma_j \| q_j - \alpha_j p_j \| ,$$

where

$$(2.41) \qquad \gamma_j = \frac{\beta_1 \tilde{d}_j' p_j + \beta_2 (\tilde{d}_j' p_j)^2 / \rho_j g_j' p_j}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j} .$$

This shows that $\tilde{d}_j' p_j > 0$ and $\beta_1 \beta_2 \geq 0$, $\beta_1 + \beta_2 \neq 0$ is a sufficient condition for $\omega_j > 0$, which by (2.37) implies that H_{j+1} is positive definite. If $\beta_1 \beta_2 < 0$, then γ_j could be zero or negative. In this case an adjustment of the parameters β_1 and β_2 is required in order to obtain a positive definite matrix H_{j+1} .

For later reference we prove the following lemma

Lemma 1

i) For every $x_j \in E^n$,

$$\| g_j \| \leq \eta \| x_j - z \| \quad \text{and} \quad \mu \| x_j - z \| \leq \| g_j \| .$$

ii) Let d_j and p_j be defined by (2.12). Then

$$d_j' p_j \geq \mu \quad \text{and} \quad \| d_j \| \leq \eta \ .$$

Proof:

i) By Taylor's theorem there are vectors

$$z_j, \hat{z}_j \in \{ x \mid x = x_j - \lambda(x_j - z) \ , \quad 0 \leq \lambda \leq 1 \}$$

such that

$$\mu \| x_j - z \|^2 \leq (x_j - z)' G(\hat{z}_j)(x_j - z) = g_j'(x_j - z) \leq \| g_j \| \ \| x_j - z \|$$

and

$$\| g_j \|^2 = g_j' G(z_j)(x_j - z) \leq \eta \| g_j \| \ \| x_j - z \| \ .$$

ii) Using Taylor's theorem again we obtain

$$d_j' p_j = \frac{(g_j - g_{j+1})' p_j}{\| \sigma_j s_j \|} = p_j' G(\hat{z}_j) p_j \geq \mu$$

and

$$\| d_j \|^2 = \frac{(g_j - g_{j+1})' d_j}{\| \sigma_j s_j \|} = d_j' G(z_j) p_j \leq \| G(z_j) \| \ \| d_j \| \leq \eta \| d_j \| \ .$$

3. Convergence

Based on the discussion of the previous section we describe now an algorithm which for any starting point x_0 and any symmetric positive definite matrix H_0 generates a sequence $\{ x_j \}$ which either terminates with z or converges superlinear to z .

At the beginning of a general cycle of the algorithm, x_j , $g_j \neq 0$ and a symmetric positive definite matrix H_j are available.

Algorithm

Step 1: (Computation of the search direction)

Compute

$$s_j = H_j g_j$$

and go to Step 2.

Step 2: (Computation of the step size σ_j)

Determine σ_j such that

$$F(x_j - \sigma_j s_j) = \min_{\sigma} \{ F(x_j - \sigma s_j) \mid \sigma \geq 0 \} .$$

Set

$$x_{j+1} = x_j - \sigma_j s_j$$

and compute g_{j+1}. If $g_{j+1} = 0$ stop, otherwise go to Step 3.

Step 3: (Computation of H_{j+1})

Select $0 \leq t \leq 1$, β_1 and β_2 with $\beta_1^2 + \beta_2^2 > 0$ and compute

$$\nu_j = F(x_{j+1}) - F(x_j) + \frac{1}{2} (g_j + g_{j+1})'(x_j - x_{j+1})$$

$$\delta_j(t) = \frac{2(t-1)\nu_j}{2t + (1-t) \| x_j - x_{j+1} \|^2}$$

$$\tilde{d}_j = \frac{g_j - g_{j+1}}{\| \sigma_j s_j \|} + \delta_j(t) \frac{s_j}{\| s_j \|} .$$

Determine H_{j+1} by formula (2.31). Replace j with $j+1$ and go to Step 1.

Remark

In Step 2 we assume that σ_j is the optimal step size. This assumption simplifies the proof that the sequence $\{x_j\}$ converges superlinearly to $\{x_j\}$. Using an approach similar to the one in [6], it can be shown that this

result remains true if σ_j is an appropriate approximation of the optimal step size. In Step 3 we assume that, if necessary, the parameters t, β_1 and β_2 are adjusted in such a way that $\tilde{d}_j'p_j = d_j'p_j + \delta_j(t) > 0$ and γ_j as defined by (2.41) is positive. As we have seen in the previous section this implies that H_{j+1} is positive definite.

In the following lemma we establish some properties of the sequences $\{g_j\}$ and $\{x_j\}$ generated by the algorithm.

Lemma 2

i) $F(x_{j+1}) \leq F(x_j) - \frac{1}{2\eta}(g_j'p_j)^2$

ii) $g_j'p_j \to 0$ as $j \to \infty$

iii) $\| x_{j+1} - x_j \| \to 0$ as $j \to \infty$.

Proof:

i) By Taylor's theorem we have, for every $\sigma \geq 0$, the inequality

$$F(x_j - \sigma s_j) \leq F(x_j) - \sigma g_j's_j + \frac{1}{2}\sigma^2 \| s_j \|^2 \eta .$$

For $\hat{\sigma}_j = g_j's_j / \|s_j\|^2 \eta$ this implies

$$F(x_j - \sigma_j s_j) \leq F(x_j - \hat{\sigma}_j s_j) \leq F(x_j) - (g_j p_j)^2 / 2\eta .$$

ii) Because $F(x)$ is bounded from below and $F(x_{j+1}) < F(x_j)$ for every j, it follows from part i) that $g_j'p_j \to 0$ as $j \to \infty$.

iii) By Taylor's theorem there is a vector z_j such that

$$z_j \in \{ x \mid x = x_j - \lambda(x_{j+1} - x_j),\ 0 \leq \lambda \leq 1 \}$$

and

$$0 = g_{j+1}'p_j = g_j'p_j - \sigma_j p_j'G(z_j)s_j .$$

Therefore,

$$(3.1) \qquad \| x_{j+1} - x_j \| = \| \sigma_j s_j \| = \frac{g_j' p_j}{p_j' G(z_j) p_j} \leq \frac{g_j' p_j}{\mu} \to 0 \text{ as } j \to \infty .$$

In order to prove that the sequence $\{x_j\}$ converges to z we need some properties of the trace of the matrices H_j and B_j. By definition, the trace of a matrix is the sum of its diagonal elements, which in turn is equal to the sum of the eigenvalues of the matrix. Using the representation (2.35) of H_j we obtain

$$(3.2) \qquad tr(H_j) = \frac{1}{\rho_j g_j' p_j} + \frac{1}{w_j' q_j} + tr(\hat{H}_j) .$$

Using (2.35) once more it is not difficult to verify that $B_j = H_j^{-1}$ can be written in the form

$$B_j = \frac{\rho_j g_j g_j'}{g_j' p_j} + \frac{w_j w_j'}{w_j' q_j} + \hat{B}_j$$

where $\hat{B}_j$ is a matrix of rank $n-2$ with the property

$$\hat{B}_j p_j = \hat{B}_j q_j = 0 .$$

Thus we have

$$(3.3) \qquad tr(B_j) = \frac{\rho_j \| g_j \|^2}{g_j' p_j} + \frac{\| w_j \|^2}{w_j' q_j} + tr(\hat{B}_j) .$$

Now let

$$\varphi_j = tr(B_j) + tr(H_j) .$$

Then it follows from (3.2), (3.3), (2.36) and

$$B_{j+1} = \frac{\tilde{d}_j \tilde{d}_j}{\tilde{d}_j' p_j} + \frac{1}{\omega_j} \frac{\| w_j \|^2}{w_j' u_j} + tr(\hat{B}_j)$$

that

$$\varphi_{j+1} = \varphi_j - \frac{1 + \| \rho_j g_j \|^2}{\rho_j g_j' p_j} + \frac{1 + \| d_j \|^2}{\tilde{d}_j' p_j} - \frac{1 + \| w_j \|^2}{w_j' q_j} + \frac{\omega_j + \| w_j \|^2 / \omega_j}{w_j' u_j} .$$

Since (2.38) and $w_j' p_j = 0$ imply

$$w_j' u_j = \frac{w_j' q_j}{\| q_j - \alpha_j p_j \|} \tag{3.4}$$

we can write the above equality in the form

$$\varphi_{j+1} = \varphi_j - \frac{1 + \| \rho_j g_j \|^2}{\rho_j g_j' p_j} + \frac{1 + \| \tilde{d}_j \|^2}{\tilde{d}_j' p_j} + \left(\frac{1}{\gamma_j} - 1 \right) \frac{\| w_j \|^2}{w_j' q_j} + \frac{\| q_j - \alpha_j p_j \|^2 \gamma_j - 1}{w_j' q_j} . \tag{3.5}$$

By Lemma 2, $\| x_{j+1} - x_j \| \to 0$ as $j \to \infty$. Therefore it follows from (2.29) and (2.30) that, for $0 \leq t \leq 1$,

$$\| \tilde{d}_j - d_j \| = | \delta_j(t) | \to 0 \quad \text{as} \quad j \to \infty .$$

In connection with Lemma 1 this implies that there is a constant μ_0 such that

$$\frac{1 + \| \tilde{d}_j \|^2}{\tilde{d}_j' p_j} \leq \mu_0 \quad \text{for } j = 0,1,\ldots . \tag{3.6}$$

Next we infer from $\tilde{d}_j = (g_j - g_{j+1}) / \| \sigma_j s_j \| + \delta_j p_j$ that $g_{j+1} - \| \sigma_j s_j \| \delta_j p_j \in S$. Defining

$$y_j = \delta_j \| \sigma_j s_j \| p_j - \frac{\| \sigma_j s_j \| \delta_j}{\tilde{d}_j' p_j} \tilde{d}_j \tag{3.7}$$

and observing that $p_j'(g_{j+1} - y_j) = 0$ we deduce from the definition of u_j and w_j that

$$u_j = \pm \frac{H_{j+1}(g_{j+1} - y_j)}{\| H_{j+1}(g_{j+1} - y_j) \|} \qquad w_j = \pm \omega_j \frac{g_{j+1} - y_j}{\| H_{j+1}(g_{j+1} - y_j) \|} . \tag{3.8}$$

Using (3.4) and (3.8) we obtain

$$(3.9) \qquad \left(\frac{1}{\gamma_j} - 1\right) \frac{\| w_j \|^2}{w_j' q_j} = (1 - \gamma_j) \frac{\| g_{j+1} - y_j \|^2}{(g_{j+1} - y_j)' H_{j+1} (g_{j+1} - y_j)}$$

and

$$(3.10) \qquad \frac{\tau_j - 1}{w_j' q_j} = \frac{\tau_j - 1}{\tau_j} \frac{\| H_{j+1}(g_{j+1} - y_j) \|^2}{(g_{j+1} - y_j)' H_{j+1} (g_{j+1} - y_j)} ,$$

where

$$(3.11) \qquad \tau_j = \gamma_j \| q_j - \alpha_j p_j \|^2 .$$

Unfortunately it does not appear to be possible to find an a priori upper bound for the expressions on the right hand side of (3.9) and (3.10). However, since it follows from (2.30) and (3.7) that $y_j = 0$ if we choose $t = 1$, there is $t_j < 1$ such that for every t with $t_j \leq t \leq 1$ the following condition is satisfied.

Condition 1

For every j the parameter t is determined such that $0 \leq t \leq 1$ and

$$(1 - \gamma_j) \frac{\| g_{j+1} - y_j \|^2}{(g_{j+1} - y_j)' H_{j+1} (g_{j+1} - y_j)} - \frac{\| g_{j+1} \|^2}{g_{j+1}' H_{j+1} g_{j+1}} \leq \mu_1$$

$$\frac{\tau_j - 1}{\tau_j} \frac{\| H_{j+1}(g_{j+1} - y_j) \|^2}{(g_{j+1} - y_j)' H_{j+1} (g_{j+1} - y_j)} - \frac{\| H_{j+1} g_{j+1} \|^2}{g_{j+1}' H_{j+1} g_{j+1}} \leq \mu_2$$

where μ_1 and μ_2 are arbitrary positive constants.

Using Condition 1 we obtain the following lemma

Lemma 3

Let Condition 1 be satisfied. Then there is a constant $\mu_3 > 0$ such that

$$\sum_{i=0}^{j} \frac{\| g_i \|}{g_i' p_i} \leq (j+1)\,\mu_3 \quad \text{for} \quad j = 0,1,2,\ldots \; .$$

Proof:

For every j we define the matrices $B_j(\gamma_j)$ and $H_j(\tau_j)$ as follows

$$B_j(\gamma_j) = B_j - (1-\gamma_{j-1}) \frac{\rho_j g_j g_j'}{g_j' p_j} \tag{3.12}$$

$$= \gamma_{j-1} \frac{\rho_j g_j g_j'}{g_j' p_j} + \frac{w_j w_j'}{w_j' q_j} + \hat{B}_j \; ,$$

$$H_j(\tau_j) = H_j - \left(1 - \frac{1}{\tau_{j-1}}\right) \frac{p_j p_j'}{\rho_j g_j' p_j} \tag{3.13}$$

$$= \frac{1}{\tau_{j-1}} \frac{p_j p_j'}{\rho_j g_j' p_j} + \frac{q_j q_j'}{w_j' q_j} + \hat{H}_j \; .$$

Setting

$$\psi_j = tr(B_j(\gamma_j) + tr(H_j(\tau_j))$$

we obtain, similar to (3.5), the equality

$$\psi_{j+1} = \psi_j - \frac{1}{\rho_j g_j' p_j}\left(\frac{1}{\tau_{j-1}} + \gamma_{j-1} \| \rho_j g_j \|^2\right) + \frac{1 + \| \tilde{d}_j \|^2}{\tilde{d}_j' p_j} \tag{3.14}$$

$$+ \left(\frac{1}{\gamma_j} - 1\right)\frac{\| w_j \|^2}{w_j' q_j} - (1-\gamma_j)\,\frac{\| \rho_{j+1} g_{j+1} \|^2}{\rho_{j+1} g_{j+1}' p_{j+1}}$$

$$+ \frac{\tau_j - 1}{w_j' q_j} - \left(1 - \frac{1}{\tau_j}\right) \frac{1}{\rho_{j+1} g_{j+1}' p_{j+1}} \; .$$

Using (3.6), (3.9) and (3.10) and observing that

$$\rho_{j+1} = \frac{1}{\| H_{j+1} g_{j+1} \|} , \quad p_{j+1} = \rho_{j+1} H_{j+1} g_{j+1}$$

we deduce from (3.14) and Condition 1 the inequality

$$\text{(3.15)} \quad \psi_{j+1} \le \psi_j - \frac{1}{\rho_j g_j' p_j} \left(\frac{1}{\tau_{j-1}} + \gamma_{j-1} \| \rho_j g_j \|^2 \right) + \mu_0 + \mu_1 + \mu_2$$

$$\le \psi_0 - \sum_{i=1}^{j} \frac{1}{\rho_i g_i' p_i} \left(\frac{1}{\tau_{j-1}} + \gamma_{i-1} \| \rho_i g_i \|^2 \right) + (j+1)(\mu_0 + \mu_1 + \mu_2) .$$

Since γ_j and τ_j are positive, $B_j(\gamma_j)$ and $H_j(\tau_j)$ are positive definite. Therefore, every ψ_j is positive and it follows from (3.15) that

$$\text{(3.16)} \quad \sum_{i=0}^{j} \frac{1}{\rho_i g_i' p_i} \left(\frac{1}{\tau_{i-1}} + \gamma_{i-1} \| \rho_i g_i \|^2 \right) \le (j+1) \mu_4 , \; j = 0,1,2,\dots ,$$

for some constant $\mu_4 > 0$.

Because $\gamma_{i-1} > 0$ and $\tau_{i-1} = \gamma_{i-1} \| q_{i-1} - \alpha_{i-1} p_{i-1} \|^2$ it is easy to verify that

$$\text{(3.17)} \quad \frac{1}{\rho_i g_i' p_i} \left(\frac{1}{\tau_{i-1}} + \gamma_{i-1} \| \rho_i g_i \|^2 \right) \ge \frac{2 \| g_i \|}{g_i' p_i} \| q_{i-1} - \alpha_{i-1} p_{i-1} \|^{-1} .$$

Moreover, for every j ,

$$\| \tilde{d}_j \| \le \eta + |\delta_j(t)| \text{ and } \tilde{d}_j' p_1 = d_j' p_j + \delta_j(t) \ge \mu + \delta_j(t) .$$

Since by (2.29), (2.30) and Lemma 2, $\delta_j(t) \to 0$ as $j \to \infty$, we have $\tilde{d}_j p_j \ge \mu/2$ for j sufficiently large. Therefore, it follows that there is $\varepsilon > 0$ such that

$$\text{(3.18)} \quad \| q_j - \alpha_j p_j \|^{-1} \ge \varepsilon > 0 \quad \text{for } j = 0,1,2,\dots .$$

The statement of the lemma follows now from (3.16), (3.17) and (3.18).

With Lemma 3 at our disposal we can now easily prove the main result of this section

Theorem 1

Let Assumption 1 and Condition 1 be satisfied. Then the sequence $\{x_j\}$ generated by the algorithm either terminates with z after a finite number of iterations or converges to z, where z is the global minimizer of $F(x)$.

Proof:

It follows from Lemma 3 that there is a positive constant μ_5 and an infinite set $J \subset \{0,1,2,...\}$ such that

$$\frac{\| g_j \|}{g_j' p_j} \leq \mu_5 \quad \text{for } j \in J .$$

Since by part ii) of Lemma 2 $p_j' g_j \to 0$ as $j \to \infty$ this inequality shows that

(3.19) $$\| g_j \| \to 0 \text{ as } j \to \infty , \quad j \in J .$$

By the continuity of $\nabla F(x)$ this implies that $\nabla F(x)$ vanishes at every cluster point of the subsequence $\{x_j, j \in J\}$. Since the set $\{x \mid F(x) \leq F(x_0)\}$ is compact and $\nabla F(x) \neq 0$ for $x \neq z$ we deduce, therefore, from (3.19) that a subsequence of $\{x_j\}$ converges to z. Observing that $F(x_{j+1}) < F(x_j)$ and $F(z) < F(x)$ for $x \neq z$, we see that $x_j \to z$ as $j \to \infty$.

4. Superlinear convergence

In order to show that the sequence $\{x_j\}$ converges superlinearly to z we need the following additional assumption.

Assumption 2

There is a positive constant L such that

$$\| G(x) - G(y) \| \leq L \| x - y \| , \quad x, y \in E^n .$$

Let $G^{1/2}$ denote the square root of $G = G(z)$ and set $G^{-1/2} = (G^{1/2})^{-1}$. As a first result we will show that with an appropriate modification of Condition 1 the sequence $\{\xi_j\}$,

$$\xi_j = tr(G^{-1/2}B(\gamma_j)G^{-1/2}) + tr(G^{1/2}H_j(\gamma_j)G^{1/2})$$

is bounded, where the matrices $B(\gamma_j)$ and $H_j(\gamma_j)$ are defined by (3.12) and (3.13), respectively.

Observing that by (2.28), (2.40) and (3.4)

$$\omega_j \frac{u_j'Gu_j}{w_j'u_j} = (\gamma_j - 1)\| q_j - \alpha_j p_j \|^2 \frac{u_j'Gu_j}{w_j'q_j} + \frac{(q_j - \alpha_j p_j)'G(q_j - \alpha_j p_j)}{w_j'q_j}$$

it is not difficult to show that in analogy to (3.14) we have

$$(4.1) \qquad \xi_{j+1} = \xi_j - \frac{1}{\rho_j g_j' p_j}\left(\frac{p_j'Gp_j}{\gamma_{j-1}} + \gamma_{j-1}\rho_j^2 g_j'G^{-1}g_j\right) + \frac{p_j'Gp_j + \tilde{d}_j'G^{-1}\tilde{d}_j}{\tilde{d}_j'p_j}$$
$$+ \left(\frac{1}{\gamma_j} - 1\right)\frac{w_j'G^{-1}w_j}{w_j'q_j} - (1-\gamma_j)\frac{\rho_{j+1}g_{j+1}'G^{-1}g_{j+1}}{g_{j+1}'p_{j+1}}$$
$$+ (\gamma_j - 1)\| q_j - \alpha_j p_j \|^2 \frac{u_j'Gu_j}{w_j'q_j} - \left(1 - \frac{1}{\gamma_j}\right)\frac{p_{j+1}'Gp_{j+1}}{\rho_{j+1}g_{j+1}'p_{j+1}}$$
$$+ \frac{(q_j - \alpha_j p_j)'G(q_j - \alpha_j p_j)}{w_j'q_j} - \frac{q_j'Gq_j}{w_j'q_j} .$$

In deriving estimates for the terms on the right hand side of the above equality we will use a result, derived in [6], which states that for every $x, y \in E^n$ with $y'x > 0$ the following relation holds

$$(4.2) \qquad \frac{x'Gx + y'G^{-1}y}{y'x} = 2 + \frac{(y - Gx)'G^{-1}(y - Gx)}{y'x}$$

Lemma 4

$$\frac{p_j'Gp_j + \tilde{d}_jG^{-1}\tilde{d}_j}{\tilde{d}_j'p_j} - 2 = O(\| x_j - z\|^2) .$$

Proof:

It follows from (2.3) and (2.4) that

$$(4.3) \qquad d_j = \frac{g_j - g_{j+1}}{\| \sigma_j s_j \|} = Gp_j + E_j p_j$$

with

$$(4.4) \qquad \| E_j \| = \| \int_0^1 G(x_j - t\sigma_j s_j)dt - G \|$$

$$\leq \| \int_0^1 G(x_j - t\sigma_j s_j)dt - G_j \| + \| G_j - G \|$$

$$\leq \max_{0 \leq t \leq 1} \{ \| G(x_j - t\sigma_j s_j) - G_j \| + \| G_j - G \| \}$$

$$\leq L \| x_{j+1} - x_j \| + L \| x_j - z \|$$

$$\leq L \| x_{j+1} - z \| + 2L \| x_j - z \|$$

$$= O(\| x_j - z \|) ,$$

where the last relation follows from (2.5) and part i) of Lemma 1.

Since by (2.29) and (2.30)

$$(4.5) \qquad |\delta_j(t)| = \leq \delta(1 - t) L \| x_{j+1} - x_j \| = O(\| x_j - z \|)$$

we infer from (4.3) and (4.4) that

$$(4.6)\qquad \| \tilde{d}_j - Gp_j \| = \| d_j + \delta_j(t)p_j - Gp_j \|$$

$$\leq \| d_j - Gp_j \| + |\delta_j(t)|$$

$$= O(\| x_j - z \|) \ .$$

Finally observing that by (4.5) and Lemma 1

$$(4.7)\qquad \tilde{d}_j'p_j = d_j'p_j + \delta_j(t) \geq \frac{1}{2}\mu \quad \text{for } j \text{ sufficiently large}$$

we obtain the statement of the lemma from (4.2) and (4.6).

In order to obtain estimates for the remaining terms on the right hand side of (4.1) we observe that by (3.8)

$$(4.8)\qquad \left(\frac{1}{\gamma_j} - 1\right) \frac{w_j'G^{-1}w_j}{w_j'q_j} = (1-\gamma_j)\frac{(g_{j+1}-y_j)'G^{-1}(g_{j+1}-y_j)}{(g_{j+1}-y_j)'H_{j+1}(g_{j+1}-y_j)}$$

and

$$(4.9)\qquad (\gamma_j - 1)\frac{\|q_j - \alpha_j p_j\|^2 u_j'Gu_j}{w_j'q_j} = \left(1-\frac{1}{\gamma_j}\right)\frac{(g_{j+1}-y_j)'H_{j+1}GH_{j+1}(g_{j+1}-y_j)}{(g_{j+1}-y_j)'H_{j+1}(g_{j+1}-y_j)} \ .$$

Because no a priori upper bound for these terms is available we introduce the following restriction on the choice of the parameter t .

<u>Condition 2</u>

For every j the parameter t is determined such that

$$|1-\gamma_j| \left| \frac{\|g_{j+1}\|^2}{(g_{j+1}-y_j)'H_{j+1}(g_{j+1}-y_j)} - \frac{\|g_{j+1}\|^2}{g_{j+1}'H_{j+1}g_{j+1}} \right| \leq \mu_6 \|g_j\|$$

$$\left|\frac{\gamma_j - 1}{\gamma_j}\right| \left| \frac{\|s_{j+1}\|^2}{(g_{j+1}-y_j)'H_{j+1}(g_{j+1}-y_j)} - \frac{\|s_{j+1}\|^2}{g_{j+1}'H_{j+1}g_{j+1}} \right| \leq \mu_6 \|g_j\|$$

$$|1-\gamma_j| \frac{\| y_j \| (\| y_j \| + \| g_{j+1} \|)}{(g_{j+1}-y_j)' H_{j+1} (g_{j+1}-y_j)} \le \mu_6 \| g_j \|$$

$$\frac{|\gamma_j - 1|}{\gamma_j} \frac{\| H_{j+1} y_j \| (\| H_{j+1} y_j \| + \| s_{j+1} \|)}{(g_{j+1}-y_j)' H_{j+1} (g_{j+1}-y_j)} \le \mu_6 \| g_j \| ,$$

where μ_6 is an arbitrary positive constant.

Since it follows from (3.7) and (2.30) that $y_j = 0$ if $t = 1$, there is $t_j < 1$ such that every $t_j \le t \le 1$ satisfies Condition 2.

Lemma 5

If Condition 2 is satisfied, then

$$\left(\frac{1}{\gamma_j} - 1 \right) \frac{w_j' G^{-1} w_j}{w_j' q_j} - (1-\gamma_j) \frac{\rho_{j+1} g_{j+1}' G^{-1} g_{j+1}}{g_{j+1}' p_{j+1}} = O(\| x_j - z \|)$$

and

$$(\gamma_j - 1) \| q_j - \alpha_j p_j \|^2 \frac{u_j' G u_j}{w_j' q_j} - \left(1 - \frac{1}{\gamma_j}\right) \frac{p_{j+1}' G p_{j+1}}{\rho_{j+1} g_{j+1}' p_{j+1}} = O(\| x_j - z \|) .$$

Proof:

Observing that $p_{j+1} = \rho_{j+1} s_{j+1}$ we obtain

$$\frac{(g_{j+1}-y_j)' G^{-1} (g_{j+1}-y_j)}{(g_{j+1}-y_j) H_{j+1} (g_{j+1}-y_j)} - \frac{g_{j+1}' G^{-1} g_{j+1}}{g_{j+1}' s_{j+1}} =$$

$$\frac{g_{j+1}' G^{-1} g_{j+1}}{(g_{j+1}-y_j)' H_{j+1} (g_{j+1}-y_j)} - \frac{g_{j+1}' G^{-1} g_{j+1}}{g_{j+1}' H_{j+1} g_{j+1}} + \frac{y_j' G^{-1} y_j - 2 y_j' G^{-1} g_{j+1}}{(g_{j+1}-y_j)' H_{j+1} (g_{j+1}-y_j)} .$$

By (4.8), Condition 2 and part i) of Lemma 1 this implies the first statement of the lemma. Similarly the second statement follows from (4.9), the relation

$$\frac{(g_{j+1}-y_j)'H_{j+1}GH_{j+1}(g_{j+1}-y_j)}{(g_{j+1}-y_j)'H_{j+1}(g_{j+1}-y_j)} - \frac{p'_{j+1}Gp_{j+1}}{\rho_{j+1}g'_{j+1}p_{j+1}} =$$

$$\frac{s'_{j+1}Gs_{j+1}}{(g_{j+1}-y_j)H_{j+1}(g_{j+1}-y_j)} - \frac{s'_{j+1}Gs_{j+1}}{g'_{j+1}s_{j+1}} + \frac{y'_jH_{j+1}GH_{j+1}y_j - 2s'_{j+1}GH_{j+1}y_j}{(g_{j+1}-y_j)'H_{j+1}(g_{j+1}-y_j)}$$

Condition 2 and part i) of Lemma 1.

The next lemma shows that the sequence $\{x_j\}$ converges sufficiently fast to z to make the sum over all numbers $\| x_j - z \|$ finite.

Lemma 6

Let Condition 1 be satisfied. Then the sum

$$\sum_{j=0}^{\infty} \| x_j - z \| \tag{4.10}$$

is finite.

Proof:

By Taylor's theorem there is a vector

$$z_j \in \{ x \mid z + \lambda(x_j - z), \;\; 0 \leq \lambda \leq 1 \}$$

such that

$$F(x_j) = F(z) + \frac{1}{2}(x_j - z)'G(z_j)(x_j - z) .$$

From this equality we deduce the relation

$$\mu \| x_j - z \|^2 \leq 2(F(x_j) - F(z)) \leq \eta \| x_j - z \|^2 . \tag{4.11}$$

Furthermore, it follows from Lemma 2 that

$$F(x_{j+1}) \leq F(x_j) - \frac{(g'_jp_j)^2}{2\eta} .$$

Subtracting $F(z)$ on both sides of this inequality and using (4.11) and Lemma 1 we obtain

$$F(x_{j+1}) - F(z) \leq F(x_j) - F(z) - \frac{(g_j'p_j)^2}{2\eta} \tag{4.12}$$

$$= \Big(F(x_j) - F(z)\Big)\left(1 - \frac{(g_j'p_j)^2}{2(F(x_j) - F(z))\eta}\right)$$

$$\leq \Big(F(x_j) - F(z)\Big)\left(1 - \frac{(g_j'p_j)^2}{\eta^2 \| x_j - z \|^2}\right)$$

$$\leq \Big(F(x_j) - F(z)\Big)\left(1 - \frac{\mu^2}{\eta^2}\, \frac{(g_j'p_j)^2}{\| g_j \|^2}\right) .$$

With

$$\zeta_j \equiv 1 - \left(\frac{\mu g_j' p_j}{\eta \, \| g_j \|}\right) \tag{4.13}$$

we infer from (4.12) that

$$F(x_{j+1}) - F(z) \leq (F(x_o) - F(z)) \prod_{i=0}^{j} \zeta_i \ . \tag{4.14}$$

For every j let $j(k)$ denote the number of elements i in the set $\{0,1,\ldots,j\}$ for which the inequality

$$\frac{\| g_i \|}{g_i' p_i} \geq 2\, \mu_3 \tag{4.15}$$

holds. Then it follows from Lemma 3 that

$$2\, \mu_3\, j(k) \leq \sum_{i=0}^{j} \frac{\| g_i \|}{g_i' p_i} \leq (j+1)\, \mu_3$$

or

$$j(k) \leq \frac{1}{2}\, (j+1) \ . \tag{4.16}$$

If we set

$$\zeta^2 = 1 - \left(\frac{\mu}{2\, \mu_3\, \eta}\right)^2$$

then (4.13), (4.15) and (4.16) imply that at least one half of the numbers ζ_i, $i = 0,1,\ldots,j$, are less than or equal to ζ^2 .

Therefore,

$$(4.17) \qquad \prod_{i=0}^{j} \zeta_i \leq \zeta^{j+1}$$

Combining (4.11), (4.14) and (4.17) we see that

$$\| x_j - z \|^2 \leq \frac{2}{\mu} (F(x_j) - F(z)) \leq \zeta^{j+1} (F(x_o) - F(z)) \frac{2}{\mu} .$$

This shows that

$$\| x_j - z \| = O(\lambda^j) \quad \text{for some} \quad 0 < \lambda < 1$$

from which it follows that the sum (4.10) is finite.

In the next lemma we establish the boundedness of the sequences $\{ \xi_j \}$, $\{ B_j(\gamma_j) \}$ and $\{ H_j(\gamma_j) \}$.

Lemma 7

Let Assumptions 1 and 2 and Conditions 1 and 2 be satisfied. Then the following statements hold

i) The sequence $\{ \xi_j \}$ is bounded.

ii) The sequences $\{B_j(\gamma_j)\}$ and $\{H_j(\gamma_j)\}$ are bounded.

iii) $\| \gamma_{j-1} \rho_j g_j - Gp_j \| \to 0$ as $j \to \infty$.

Proof:

i) It follows from the equality (4.1) and Lemmas 4 and 5 that there is a constant μ_7 such that

$$(4.18) \qquad \xi_{j+1} \leq \xi_j + \mu_7 \| x_j - z \| + \frac{\alpha_j^2 \, p_j' G p_j - 2\alpha_j p_j' G q_j}{w_j' q_j} .$$

Since $\alpha_j = \tilde{d}_j' q_j / \tilde{d}_j' p_j$ and by (3.6) $\tilde{d}_j' p_j > 0$ is bounded away from zero, we have

$$(4.19) \qquad \gamma_j^2 p_j'Gp_j - 2\gamma_j p_j'Gq_j = -\frac{(\tilde{d}_j'q_j)^2}{\tilde{d}_j'p_j} + (p_j'G - \tilde{d}_j')p_j \left(\frac{\tilde{d}_j'q_j}{\tilde{d}_j'p_j}\right)^2$$

$$- 2(p_j'G - \tilde{d}_j')\, q_j \, \frac{\tilde{d}_j'q_j}{\tilde{d}_j'p_j}$$

$$= O(\| \tilde{d}_j - Gp_j \|) = O(\| x_j - z \|) ,$$

where the last equality follows from (4.6).

Observing that by (4.2), $\xi_j \geq 1$ and $\xi_j \geq 1 / w_j'q_j$ we deduce from (4.18) and (4.19) the inequality

$$\xi_{j+1} \leq \xi_j + \mu_7 \|x_j - z\| + O\left(\frac{\| x_j - z \|}{w_j'q_j}\right)$$

$$\leq \xi_j (1 + \mu_8 \|x_j - z\|)$$

$$= \xi_0 \prod_{i=0}^{j} (1 + \mu_8 \| x_i - z \|) ,$$

where μ_8 is a suitable positive constant. Since by Lemma 6 the sum

$$\sum_{j=1}^{\infty} \| x_j - z \|$$

is finite this implies that $\xi_j \leq \mu_9$ for some constant μ_9 and all $j \geq 0$.

ii) For each j , the matrices $B_j(\gamma_j)$ and $H_j(\gamma_j)$ are symmetric and positive definite. Hence all their eigenvalues are real and positive. Since by definition, ξ_j is equal to the sum of their eigenvalues it follows from part i) of the lemma that the sequences $\{B_j(\gamma_j)\}$ and $\{H_j(\gamma_j)\}$ are bounded.

iii) By part i) of the lemma the sequence $\{1 / w_j'q_j\}$ is bounded. Therefore, (4.19) implies that

$$(4.20) \qquad \frac{(q_j - \alpha_j p_j)'G(q_j - \alpha_j p_j)}{w_j' q_j} - \frac{q_j' G q_j}{w_j' q_j} = O(\| x_j - z \|) .$$

Since α_j 0 for all j, it follows from (4.1), (4.20) and Lemmas 4, 5 and 6 that the sum

$$\sum_{j=0}^{\infty} \left(\frac{p_j' G p_j + \gamma_{j-1}^2 \rho_j^2 g_j' G^{-1} g_j}{\gamma_{j-1} \rho_j g_j' p_j} - 2 \right)$$

is finite. Hence,

$$\frac{p_j' G p_j - \gamma_{j-1}^2 \rho_j' g_j' G^{-1} g_j}{\gamma_{j-1} \rho_j g_j p_j} \to 2 \quad \text{as} \quad j \to \infty ,$$

which by (4.2) implies

$$\| \gamma_{j-1} \rho_j g_j' p_j - G p_j \| \to 0 \quad \text{as} \quad j \to \infty .$$

Using the above lemmas we can now prove the main result of this section.

Theorem 2

Let Assumptions 1 and 2 and Conditions 1 and 2 be satisfied. Then

$$\frac{\| x_{j+1} - z \|}{\| x_j - z \|} \to 0 \quad \text{as} \quad j \to \infty .$$

Proof:

It follows from (3.1) that

$$(4.21) \qquad \| \sigma_j s_j \| = \frac{g_j' p_j}{p_j' G p_j + p_j' (G(z_j) - G) p_j}$$

$$= \frac{g_j' p_j}{\gamma_{j-1} \rho_j g_j' p_j + (p_j' G - \gamma_{j-1} \rho_j g_j') p_j + p_j' (G(z_j) - G) p_j}$$

Since

$$(4.22) \qquad \| G(z_j) - G \| \leq \| G(z_j) - G_j \| + \| G_j - G \| \to 0 \quad \text{as} \quad j \to \infty$$

and by part iii) of Lemma 7

$$\| \gamma_{j-1}\rho_j g_j - Gp_j \| \to 0 \quad \text{as} \quad j \to \infty \tag{4.23}$$

and

$$\gamma_{j-1}\rho_j g_j' p_j \geq \mu/2 \quad \text{for } j \text{ sufficiently large}$$

we infer from (4.21) that

$$\| \sigma_j s_j \| \gamma_{j-1}\rho_j = \sigma_j\gamma_{j-1} \to 1 \quad \text{as} \quad j \to \infty . \tag{4.24}$$

Furthermore,

$$\left\| \frac{g_j}{\| g_j \|} - G \frac{\sigma_j s_j}{\| g_j \|} \right\| = \left\| \frac{g_j}{\| g_j \|} - G \frac{s_j}{\gamma_{j-1} \| g_j \|} - G \frac{(\sigma_j\gamma_{j-1} - 1)s_j}{\gamma_{j-1} \| g_j \|} \right\| \tag{4.25}$$

$$\leq \left\| \frac{g_j}{\| g_j \|} - G \frac{s_j}{\gamma_{j-1} \| g_j \|} \right\| + \| G \| \frac{\| s_j \|}{\gamma_{j-1} \| g_j \|} (\sigma_j\gamma_{j-1} - 1).$$

and

$$\| \gamma_{j-1}\rho_j g_j - Gp_j \| = \gamma_{j-1}\rho_j \| g_j \| \left\| \frac{g_j}{\| g_j \|} - G \frac{s_j}{\gamma_{j-1} \| g_j \|} \right\| .$$

Therefore,

$$\left\| \frac{g_j}{\| g_j \|} - G \frac{s_j}{\gamma_{j-1} \| g_j \|} \right\| = \frac{1}{\gamma_{j-1}\rho_j \| g_j \|} \left\| \gamma_{j-1}\rho_j g_j - Gp_j \right\| . \tag{4.26}$$

Observing that by part iii) of Lemma 7 the sequence

$$\left\{ \frac{1}{\gamma_j\rho_j \| g_j \|} \right\} = \left\{ \frac{\| s_j \|}{\gamma_j \| g_j \|} \right\}$$

is bounded we deduce from (4.23) through (4.26) that

$$\left\| \frac{g_j}{\| g_j \|} - G \frac{\sigma_j s_j}{\| g_j \|} \right\| \to 0 \quad \text{as} \quad j \to \infty .$$

In conjunction with (2.5), (2.7) and part i) of Lemma 1 this completes the proof of the theorem.

In the following lemma it is shown that the sequence $\{\gamma_j\}$ converges to one.

Lemma 8

Let Assumptions 1 and 2 and Conditions 1 and 2 be satisfied. For every choice of β_1 and β_2 such that $\beta_1 \beta_2 \geq 0$ and $\beta_1 + \beta_2 \neq 0$ the following statements hold.

i) $\quad 0 < \gamma_j \leq 1 \quad j = 0,1,2,\ldots$

ii) $\quad |1 - \gamma_j| = O\left(\max\left\{\frac{\|x_{j+1} - z\|^2}{\|x_j - z\|^2}, \|x_j - z\|^2\right\}\right)$.

Proof:

It follows from (2.41) and (2.39) that

$$(4.27) \qquad \gamma_j = \frac{\beta_1 \tilde{d}_j' p_j + \beta_2(\tilde{d}_j' H_j \tilde{d}_j - (\tilde{d}_j' q_j)^2 / w_j' q_j)}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j}$$

$$= 1 - \frac{\beta_2 (\tilde{d}_j' q_j)^2 / w_j' q_j}{\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j}$$

Because $\beta_2(\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j) \geq 0$ the above equality proves the first part of the lemma. Furthermore, the relation

$$|\beta_2| \, \tilde{d}_j' H_j \tilde{d}_j \leq |\beta_1 \tilde{d}_j' p_j + \beta_2 \tilde{d}_j' H_j \tilde{d}_j|$$

shows that it suffices to prove the second part of the lemma for the case $\beta_1 = 0$ and $\beta_2 = 1$. With this choice of parameters we obtain from (2.41) the equality

$$(4.28) \qquad \frac{1}{\gamma_j} = \frac{\rho_j g_j' p_j \tilde{d}_j' H_j \tilde{d}_j}{(\tilde{d}_j' p_j)^2} = \frac{(\tilde{d}_j' p_j)^2 + \frac{\rho_j g_j' p_j}{w_j' q_j} (\tilde{d}_j' q_j)^2}{(\tilde{d}_j' p_j)^2}$$

$$= 1 + \frac{\rho_j g_j' p_j}{w_j' q_j} (\tilde{d}_j' q_j)^2$$

Because $1 / w_j' q_j \leq \gamma_j$ and

$$\rho_j g_j' p_j = \frac{1}{\gamma_{j-1}} \left(p_j' G p_j + (\gamma_{j-1} \rho_j g_j p_j - p_j' G p_j) \right)$$

it follows from (4.28) and parts i) and iii) of Lemma 7 that there is some constant $\mu_{1o} > 0$ such that

$$\frac{1}{\gamma_j} \leq 1 + \frac{\mu_{1o}}{\gamma_{j-1}} (\tilde{d}_j' q_j)^2 , \quad j = 0,1,2,\ldots \quad . \tag{4.29}$$

Finally observe that

$$\begin{aligned} |\tilde{d}_j' q_j| &\leq |d_j' q_j| + |\delta_j(t) p_j' q_j| \\ &\leq \frac{|(g_j' - g_{j+1})' q_j|}{\| \sigma_j s_j \|} + |\delta_j(t)| \\ &= \frac{|g_{j+1}' q_j|}{\| \sigma_j s_j \|} + |\delta_j(t)| \quad . \end{aligned} \tag{4.30}$$

Since by (3.1) and part ii) of Lemma 7

$$\frac{1}{\|\sigma_j s_j\|} = \frac{p_j' G(z_j) p_j}{g_j' p_j} \leq \frac{\eta}{g_j' p_j} = 0 \left(\frac{1}{\| g_j \|} \right)$$

and by (2.30) and (2.29)

$$|\delta_j(t)| = 0 (\| x_{j+1} - x_j \|) = 0 (\| x_j - z \|) ,$$

it follows from (4.30) and part i) of Lemma 1 that

$$|\tilde{d}_j' q_j| = 0 \left(\frac{\| x_{j+1} - z \|}{\| x_j - z \|} + \| x_j - z \| \right) \quad . \tag{4.31}$$

By Theorem 2, the right hand side of (4.31) converges to zero which by (4.29) implies that the sequence $\{ 1 / \gamma_j \}$ is bounded. Hence we obtain from (4.29) the equality

$$\left| \frac{1}{\gamma_j} - 1 \right| = 0 \left((\tilde{d}_j' q_j)^2 \right) \quad .$$

Because $d_j'q_j \to 0$ as $j \to \infty$ this implies

$$|\gamma_j - 1| = O\left((\tilde{d}_j'q_j)^2 \right) .$$

In view of (4.31) this relation completes the proof of the lemma.

In the final theorem it is shown that, for j sufficiently large, every value of t between zero and one satisfies Conditions 1 and 2.

Theorem 3

Let Assumptions 1 and 2 and Conditions 1 and 2 be satisfied and suppose that $\beta_1 \beta_2 \geq 0$ and $\beta_1 + \beta_2 \neq 0$. Then

i) The sequences $\{B_j\}$ and $\{H_j\}$ are bounded.

ii) $\sigma_j \to 1$ as $j \to \infty$.

iii) There is j_0 such that, for $j \geq j_0$, every t , $0 \leq t \leq 1$, satisfies Conditions 1 and 2.

Proof:

Because it follows from Lemma 8 that

$$\gamma_j \to 1 \text{ as } j \to \infty \tag{4.32}$$

the first part of the theorem follows from part ii) of Lemma 7. In order to prove the second part of the theorem we deduce from (4.21) the relation

$$\sigma_j = \frac{g_j'p_j}{g_j'p_j + (\gamma_{j-1} - 1)g_j'p_j + \|s_j\| p_j'(G(z_j) - G)p_j + \|s_j\|(p_j'G - \gamma_{j-1}\rho_j g_j)'p_j}$$

which by (4.32), (4.22), part iii) of Lemma 7 and the boundedness of the sequence $\{ \|s_j\| / g_j'p_j \}$ implies that $\sigma_j \to 1$ as $j \to \infty$.

Because $\alpha_j = \tilde{d}_j'q_j / \tilde{d}_j'p_j$ it follows from (3.11) and (4.7) that

$$\tau_j = O(\gamma_j) .$$

Thus Lemma 8 and part i of Theorem 3 imply that every t, $0 \leq t \leq 1$, satisfies Condition 1 for j sufficiently large. Furthermore, using (3.7), (4.7) and the fact that $g'_{j+1}p_j = 0$, it is not difficult to verify that there is $\varepsilon > 0$ such that

$$(4.33) \quad \min\left\{ \left\| \frac{g_{j+1}}{\|g_{j+1}\|} - \frac{y_j}{\|g_{j+1}\|} \right\|, \left\| \frac{g_{j+1}}{\|y_j\|} - \frac{y_j}{\|y_j\|} \right\| \right\} \geq \varepsilon, \quad 1 = 0,1,2,\dots .$$

Because by (4.5)

$$|\delta_j(t)| = O(\|x_j - z\|), \quad 0 \leq t \leq 1,$$

it follows from (3.7) that

$$\|y_j\| = O(\|x_j - z\|^2),$$

we deduce from (4.33), Lemma 1, and the first part of the theorem the relation

$$(4.34) \quad \frac{\|g_{j+1}\|^2}{(g_{j+1} - y_j)'H_{j+1}(g_{j+1} - y_j)} - \frac{\|g_{j+1}\|^2}{g'_{j+1}H_{j+1}g_{j+1}} = O\left(\min\left\{1, \frac{\|x_j - z\|^2}{\|x_{j+1} - z\|}\right\}\right).$$

Since by Lemma 8

$$|1 - \gamma_j| = O \max\left\{\|x_j - z\|^2, \frac{\|x_{j+1} - z\|^2}{\|x_j - z\|^2}\right\}$$

we obtain from (4.34) the equality

$$|1 - \gamma_j| \left| \frac{\|g_{j+1}\|^2}{(g_{j+1} - y_j)'H_{j+1}(g_{j+1} - y_j)} - \frac{\|g_{j+1}\|^2}{g'_{j+1}H_{j+1}g_{j+1}} \right| = O\left(\max\{\|x_j - z\|^2, \|x_{j+1} - z\|\}\right).$$

Using Lemma 1 and Theorem 2 we see that the first inequality of Condition 2 is satisfied for all t, $0 \leq t \leq 1$, if j is sufficiently large. Since a completely analoguous argument shows that the remaining inequalities are also satisfied for all t, $0 \leq t \leq 1$, and all sufficiently large j this completes the proof of the theorem.

REFERENCES

[1] M.C. Biggs, "Minimization algorithms making use of non-quadratic properties of the objective function", Journal of the Institute of Mathematics and its Applications 8 (1971), pp. 315-327.

[2] C.G. Broyden, "Quasi-Newton methods and their application to function minimization", Mathematics of Computation 21 (1967), pp. 368-381.

[3] C.G. Broyden, "The convergence of a class of double-rank minimization algorithms", Parts 1 and 2, Journal of the Institute of Mathematics and its Applications 6 (1970), pp. 76-90, pp. 222-231.

[4] R. Fletcher, "A new approach to variable metric algorithms", The Computer Journal 13 (1970), pp. 317-322.

[5] D. Goldfarb, "A family of variable metric methods derived by variational means", Mathematics of Computation 24 (1970), pp. 23-26.

[6] K. Ritter, "Global and superlinear convergence of a class of variable metric methods", Mathematics Research Center TSR No. 1945, University of Wisconsin-Madison, 1979.

[7] R.B. Schnabel, "Q-superlinear convergence of Biggs' method and related methods for unconstrained optimization", Department of Computer Science TR No. CU-CS-133-78, University of Colorado, Boulder, 1978.

[8] D.F. Shanno, "Conditioning of quasi-Newton methods for function minimization", Mathematics of Computation 24 (1970), pp. 647-656.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #1967	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) GLOBAL AND SUPERLINEAR CONVERGENCE OF A CLASS OF SCALED VARIABLE METRIC METHODS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Klaus/Ritter		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE (11) June 1979
		13. NUMBER OF PAGES 37 (12) 41
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) MRC-TSR-1967		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(9) Technical summary rept.,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Unconstrained minimization, variable metric method, global convergence, superlinear convergence.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper considers a class of variable metric methods for unconstrained minimization. The update formulas are such that the quasi-Newton equation is not necessarily satisfied. Under appropriate assumptions on the function to be minimized, each algorithm in this class converges globally and superlinearly.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

221 200

JCB